

**SYSTEM AND METHOD FOR DATA COLLECTION, EVALUATION,  
INFORMATION GENERATION, AND PRESENTATION**

**Inventor: Georg J. Rebane, Ph.D.**  
**Assigned to: BIZRATE.COM**

5

**BACKGROUND OF THE INVENTION**

10 The present invention relates to a system and method for data collection, evaluation, information generation, and presentation. More particularly the present invention relates to a system for collecting, evaluating, and presenting data, and generating information relating to electronic commerce. The system and methods of the present invention include one or more of the following: a

15 module for stabilizing small or noisy samples of data; alarm modules that alert an event handler when data values cross specified thresholds; predictor modules that use recent historical data along with an estimated and/or available saturation population function as the basis for a differential equation that defines the growth of the population to a maximum attainable level; and a dynamic icon that conveys

20 to users of a system levels of predefined activity occurring on the system.

The availability of relatively low cost, powerful computer systems and the development of online communication systems and networks—principally the Internet and its protocols, and the availability of low-cost consumer computer systems—have fueled the growth of e-commerce. As used herein "e-commerce" 25 means commercial transactions for goods or services, particularly wholesale or

retail sales of products or services, or bartered exchange of the foregoing, over global computer networks, such as the Internet, or any smaller computer network that unites users and suppliers of goods or services.

The rapid growth of e-commerce makes the need for such guidance even 5 more compelling. In a study by the assignee of the present invention, it was found that over the 12-month period ending June 1999, total retail e-commerce sales tripled from \$2.67 to \$7.94 billion. (Source: 2<sup>ND</sup> QTR: 1999 Consumer Online Report for Total Retail e-commerce, published by BizRate.com, 1999).

The Internet has been swiftly facilitating the growth of local and regional 10 markets into national and international markets. This market expansion provides consumers with many new advantages and opportunities including better product pricing, product selection, product quality, and customer service. The market expansion also creates new advantages and opportunities for businesses, including a broader base of consumers and suppliers. With the advantages and 15 opportunities come new challenges.

The success of a business engaged in e-commerce may depend on how well it understands the dynamics and parameters of the e-commerce marketplace, and how well the business understands its status in such marketplace. Unfortunately, traditional models for evaluating the performance of 20 a business are not well suited or optimized for evaluating the performances of businesses engaged in e-commerce. The nature of e-commerce and the manner in which it is conducted demand new and improved systems and methods for

evaluating business performance. Consumers also need guidance so that they may understand their options and make the best decisions when doing business online.

From the consumer perspective, chief among the challenges is finding

5 among the myriad of online businesses the merchants who offer the category of products sought at the best pricing, product selection, product quality, reputation, etc. The rating of a merchant relative to such variables may change rapidly in response to marketplace conditions. For example, marketplace competition may drive competitors to change their prices daily. Product availability may also  
10 change on as frequent a basis. For such reasons, consumers need a mechanism that helps them quickly locate the best merchants for their needs based on the most current and accurate data and information available.

Merchants also face new challenges in the online marketplace. They must

be able to reach consumers and communicate to consumers the value they can

15 deliver. To do this, they must understand the competition, and what drives consumers to make purchases. They must monitor their own prices relative to competitors' prices on a frequent basis to remain competitive. They must also understand what level of satisfaction or dissatisfaction consumers have from transactions with themselves and competitors, as well as the bases of  
20 satisfaction or dissatisfaction. For such reasons, merchants need a mechanism that helps them quickly obtain the most current and accurate information.

Traditionally, marketing surveys have been employed to determine how well a business rates in the eyes of consumers. Direct feedback from consumers provides important information. In traditional forms of commerce, consumer satisfaction surveys have been long used to gather direct feedback from 5 consumers. (Traditional forms of commerce include in-store, telephonic, and mail order commerce.) The surveys help businesses understand what positive and negative things they are doing. With the proper understanding, the businesses may reinforce the positive things and correct the negative ones. The more accurate and current the survey results, the better a business can achieve 10 its objectives.

Unfortunately, administering and processing consumer surveys, even in traditional modes of commerce, has been problematic in various respects, for example:

- How do you get consumers to fill out such surveys?
- Is the sample size of responses large enough to produce accurate 15 reports?
- How do you input and process the data?
- Has the data been processed in a timely manner?
- How are results to be timely reported to businesses?
- What do the results mean to the business?
- What steps should the business take in view of the results?

- What has been the effect of any remedial action?

With respect to these questions, there are many problems. The surveys are typically presented to consumers as paper questionnaires for the consumers to manually fill in. Such surveys may be costly to construct and print. Once 5 printed, they cannot be modified. Consumers generally dislike filling out survey questionnaires, and therefore it may be difficult or time consuming for the survey sponsor to gather enough completed survey questionnaires to constitute a statistically significant sample size. Often, the completed survey questionnaires must be returned by mail; even if a consumer has filled out the survey 10 questionnaire, the consumer may not take the trouble of dropping it in the mailbox.

To overcome these kinds of problems, survey sponsors sometimes employ individuals to field survey responses from individuals. These individuals may be stationed in a store to verbally field answers to survey questions or they 15 may telephone consumers after a transaction. The problem with using individuals to administer surveys is the cost of administering the survey and the intrusiveness of the process. The intrusiveness is such that consumers may be alienated from doing further business with a merchant. This is particularly a concern relative to telephone surveys. With intrusive survey methods, even if the 20 consumer is inclined to answer questions by a survey taker, the consumer's answers may be skewed toward an unfavorable response, creating inaccurate results.

The processing of completed survey questionnaires has its own set of disadvantages. The completed survey questionnaires usually must be read by data entry personnel and manually input by them into a data processing system. This not only adds to the cost of administering a survey, but it also results in

5   delays between the time surveys are taken and the time the results are processed.

In view of the inherent delays in administering surveys and inputting survey data, by the time a business receives a report of the results of a survey, the results may no longer be accurate. For example, consider a survey about

10   price competitiveness: competitors may have dropped prices in the interval between the responses to a survey and the processing and reporting of information. During such an interval, a business could lose significant sales and revenues because they have not reacted contemporaneously to competitors' price changes. If there is a decline in consumer service ratings that is not

15   corrected quickly because of the interval between survey responses and reporting of results, a business may also lose consumer goodwill, and consequently sales and revenues.

Another problem with traditional survey modes is that a business may have difficulty benefiting from the survey results. One reason is that, although a

20   business can take remedial action in view of the results, to monitor the results of any remedial action requires a subsequent survey. There is a disincentive for a subsequent survey because of all the aforementioned disadvantages related to

traditional modes of surveying such as cost, time required, etc. However, if the problems inherent in traditional forms of surveying could be overcome, subsequent surveys could be undertaken to determine the effectiveness of remedial action dictated by the initial survey.

5        In view of the disadvantages in traditional survey methodology, a few years ago the assignee of the present invention developed and implemented a novel system for providing timely and accurate reporting of information relating to the sales, marketing, consumer satisfaction, and other commercial activities of participating businesses. In the system, online buyers are non-intrusively invited

10      to fill out a survey questionnaire immediately after completing a purchase at a participating merchant. The invitation is in the form of a banner on the order confirmation receipt from a participating merchant's website. If the buyer has clicked on the banner to accept the survey questionnaire, the buyer is hyper-linked to a questionnaire from a survey system server. The buyer completes the

15      survey questionnaire and the survey data are electronically returned to a data processing system for processing and evaluating survey results. The same system can electronically report the processed results from a sample of survey questionnaires directly to participating or subscribing merchants.

Thus, the system of the assignee overcomes disadvantages in the art by

20      providing a system that electronically gathers data and transmits it directly into the data processing system. Among the advantages of this system, it eliminates the need for manual entry of data on paper forms; the use of individuals to take

and input survey data gathered from consumers; the delays that occur between data collection and input, and data input and processing; and the costs associated with such methodologies. While this system has begun to address many disadvantages of traditional modes of surveying, processing, and

5 evaluating survey data, the dynamics of the e-commerce marketplace demand faster and more accurate data gathering, processing, evaluation and reporting of data and information.

The rate at which reports can be issued depends on how fast survey responses are returned and on the minimum sample size required. It is

10 fundamental in statistical sciences that, when conventional modes of statistical analysis are employed, an inadequately small sample or a noisy sample will lead to inaccurate results. However, business and marketplace conditions may be changing faster than adequate sample sizes can be gathered for accurate results using conventional modes of statistical analysis and estimation theory. Such

15 conventional modes include "moving window averages" (weighted and unweighted). Accordingly, businesses may be at risk if certain trends relating to the business or marketplace take shape before data is collected in sample sizes suitable for traditional modes of statistical analysis and estimation.

In view of the foregoing, there is a strong need for novel data processing systems that can discern trends and otherwise provide results based on limited or noisy data samples. Further in view of the foregoing, there is a substantial need for data gathering, processing and evaluation systems that quickly alert

businesses to incipient trends in their business activities and marketplace so that appropriate action may be taken to protect and advance a business's well-being.

There is also a need for systems that allow a business to predict growth rates and limits of variables relating to the business or marketplace.

5

### SUMMARY OF THE INVENTION

The present invention relates to a system and method for data collection, information generation, evaluation, and presentation that overcomes the aforementioned problems in the prior art. More particularly the present invention

- 10 relates to a system for collecting data, generating, evaluating, and presenting information relating to electronic commerce via the Internet. The system and methods of the present invention include one or more of the following: a module for stabilizing small or noisy samples of data; alarm modules that alert a handler when data values are anomalous or cross specified thresholds; predictor
- 15 modules that use recent historical data along with an estimated and/or available saturation population function as the basis for a differential equation that predicts the future growth of the population to a maximum attainable level; and a dynamic measurement indicator that conveys to users of a system levels of predefined and ongoing activity occurring on another system. The fields of the invention
- 20 include ecommerce; information retrieval/analysis; and planning and control.

Before the present invention, the period for gathering sufficient quantities of data

to resolve a trend often exceeded the time required for a merchant to begin suffering lost sales or other harms because of an undetected incipient trend.

In one novel embodiment, the present invention provides a system for data collection, evaluation, information generation, and/or presentation

5 comprising a data capture server capable of receiving data from a data source; one or more databases for receiving data from the data capture server; a plurality of processing modules in communication with each other and/or the one or more databases, each processing module performing a predefined operation on data stored in a database or received from a processing module, at least two

10 processing modules being selected from the group consisting of: a statistical analysis processing module; a data stabilizer processing module; a saturation limited forecasting module; a dynamic activity-level icon module; and an alarm filter module; one or more databases in communication with one or more processing modules for storing processed data received from a selected

15 processing module; and a presentation server in communication with one or more of the databases for receiving items of data stored therein and presenting selected items of data. In the foregoing embodiment, any combination of processing modules may be selected. The foregoing embodiment may provide for the data capture server to be communication with a data source comprising

20 one or more remote computer systems. The data capture server may be adapted to receive data from computer systems of consumers following an online e-commerce transaction.

In another novel embodiment, the present invention provides a system for data collection, evaluation, information generation, and/or presentation comprising a data capture server capable of receiving data from a data source over a computer network, the data source providing data related to a transaction

5 between buyers and sellers; one or more databases for receiving data from the data capture server; a plurality of processing modules in communication with each other and/or one or more of the databases, each processing module performing a predefined operation on data stored in a database or received from a processing module, at least two processing modules being selected from the

10 group consisting of: a statistical analysis processing module; a data stabilizer processing module; a saturation limited forecasting module; a dynamic activity-level icon module; and an alarm filter module; one or more databases in communication with one or more processing modules for storing processed data received from a selected processing module; and a presentation server in

15 communication with one or more of the databases for receiving items of data stored therein and presenting selected items of data as data or information, data on the presentation server being accessible to remote computer systems via a network. In the foregoing embodiment, the network over which the data source and data capture server communicate may be the Internet, and the presentation

20 server may be accessible by remote computer systems via the Internet. The foregoing embodiments may further provide for a survey server that serves a survey questionnaire to a remote computer system comprising a data source so

that a user of a remote computer system comprising the data source can complete the survey questionnaire, a completed survey questionnaire containing data supplied by the user being returnable to the data capture server over the Internet. The remote computer systems may be a plurality of consumer and/or

5 merchant computer systems. A consumer may complete a survey questionnaire with data about an online transaction between the consumer and a merchant. In the foregoing systems, the presentation server may serve data comprising ratings about online merchants, the ratings being based on data collected by the data capture server from consumer computer systems.

10 In a further novel embodiment, the present invention provides a presentation server for serving web pages to a plurality of remote computer systems over a network; the web pages including ratings information for one or more merchants; the ratings information being derived from data captured from buyer computer systems and processed by a data stabilizer processing module.

15 The web page may further include a link to the website of a rated merchant. It may also include a dynamic activity-level icon for showing the level of a predetermined activity at a rated merchant website. The web page may further include information about incentive programs offered to a user of a remote computer system.

20 In another novel embodiment, the present invention provides a presentation server for serving web pages via a network to a plurality of remote computer systems; one or more web pages including a dynamic activity-level

icon for indicating a level of a predetermined activity at one or more predetermined websites. A web page may include at least two dynamic activity-level icons, each being for competitive merchants and the web page is accessible by remote computer systems of consumers. In the foregoing

5 embodiment, the network may be the Internet. The web page in the foregoing embodiment may include a reference dynamic activity-level icon. In the foregoing embodiment, one or more dynamic activity-level icons on a web page may be each associated with a link to a corresponding merchant website. In the foregoing embodiment, the ratings for a merchant may also be associated with

10 the dynamic activity-level icon and the link to the merchant website.

In another novel embodiment, the present invention provides a presentation server for serving web pages via a network to a plurality of remote computer systems; one or more web pages including a dynamic activity-level icon for indicating a level of a predetermined activity for one or more products

15 being sold by one or more merchants over a computer network. The presentation server may make the web pages accessible to consumer computer systems via a network comprising the Internet.

In another novel embodiment, the present invention provides a presentation server that includes and serves a dynamic activity-level applet via a

20 network to a plurality of remote computer systems, the applet enabling a remote computer system to access a data source that communicates data representative of a predefined activity level, the applet implementing a dynamic activity-level

icon on a display of the remote computer system based on and responsive to the data communicated to it. The presentation server may serve the applet to remote computer systems comprising consumer computer systems. The network may be the Internet.

5        In another novel embodiment, the presentation server provides a method for data collection, evaluation, information generation, and/or presentation comprising capturing data from a data source in a data capture server; transferring data from the data capture server to one or more databases for receiving data; transferring data from the one or more databases to one or more

10      processing modules, each processing module capable of performing a predefined operation on transferred data, at least two processing modules being selected from the group consisting of: a statistical analysis processing module; a data stabilizer processing module; a saturation limited forecasting module; a dynamic activity-level icon module; and an alarm filter module, the processing

15      module outputting processed data or information; presenting selected items of processed data or information via a presentation server. The data may be transferred to the data capture server by one or more remote computer systems via a network, which may be the Internet, for example. The captured data may relate e-commerce transactions. The online transactions may be consumer-merchant transactions and/or business to business transactions. The data may be transferred to the data capture server via consumer computer systems. The method may also include providing a survey server that serves a survey

20

questionnaire to one or more remote computer systems comprising data sources so that a user of a remote computer system comprising the data source can complete the survey questionnaire, a completed survey questionnaire containing data supplied by the user being returnable to the data capture server over the

5 Internet. In the method, the one or more remote computer systems may be a plurality of consumer and/or merchant computer systems and the completed survey questionnaire contain data about an online transaction between the consumer and a merchant. In the method, the presentation server may serve data comprising ratings about online merchants. The ratings may be derived

10 from data collected by the data capture server from consumer computer systems.

In another novel embodiment, the present invention provides a method comprising deriving ratings for one or more merchants using a data stabilizer, and serving web pages to a plurality of remote networked computer systems, the web pages including the derived ratings information for one or more merchants.

15 The served web pages may include a link to the website of a rated merchant. A web page may include a dynamic activity-level icon for showing the level of a predetermined activity at a rated merchant website. The web page may include information about incentive programs offered to a user. The method may include associating one or more dynamic activity-level icons for indicating a level of a

20 predetermined activity at one or more predetermined websites with one or more web pages, and serving the one or more web pages via a network to a plurality of remote computer systems. In the method, a web page may include at least two

dynamic activity-level icons, each being for competitive merchants and the web page is accessible by remote computer systems of consumers. A reference dynamic activity-level icon may also be included. One or more dynamic activity-level icons on a web page may be associated with a link to a corresponding merchant website. The ratings for a merchant may also be associated with the dynamic activity-level icon and/or the link to the merchant website. In the method, a web page may include at least two dynamic activity-level icons, one indicating the level of a predefined activity for a merchant website, the other being a reference icon for indicating a predetermined level of the predefined activity, the web page being accessible by a consumer remote computer system.

In another embodiment, the present invention provides a method comprising serving web pages via a network comprising the Internet to a plurality of remote consumer computer systems; one or more web pages including a dynamic activity-level icon for indicating a level of a predetermined activity for one or more products being sold by one or more merchants over a computer network. A web page may include at least two dynamic activity-level icons, each being for competitive merchants and the web page is accessible by remote computer systems of consumers. The dynamic activity-level icons may relate to relative buying activity at the merchant websites.

In another novel embodiment, the present invention provides a method comprising serving a dynamic activity-level applet via a network to a plurality of remote computer systems, the applet enabling a remote computer system to

access a data source that communicates data representative of a predefined activity level, the applet implementing one or more dynamic activity-level icons on a display of the remote computer system based on and responsive to the data communicated to it. The presentation server serves the applet to remote

- 5 computer systems comprising consumer computer systems. In the method, data is communicated to a consumer computer system with the applet for input into the applet, the applet implementing the dynamic activity-level icon in accordance to the input data. The communicated data may relate to activities at one or more merchant websites. For example, the communicated data may relate to relative
- 10 buying activities at competitive merchant websites. Alternatively, for example, the communicated data may relate to activity levels for one or more of a product, product-type, and product category.

In another novel embodiment, the present invention provides a presentation server that includes web pages containing data or information that

- 15 has been derived from at least two processing modules selected from the group consisting of a statistical analysis processing module; a data stabilizer processing module; a saturation limited forecasting module; a dynamic activity-level icon module; and an alarm filter module, the web pages being accessible to a plurality of remote merchant systems over a computer network, such as the

- 20 Internet.

In another novel embodiment, the present invention provides a presentation server that includes web pages containing data or information that

has been derived from at least two processing modules selected from the group consisting of a statistical analysis processing module; a data stabilizer processing module; a saturation limited forecasting module; a dynamic activity-level icon module; and an alarm filter module, the web pages being accessible to

5 a plurality of remote consumer computer systems over a computer network, such as the Internet. In the foregoing embodiments, the web pages may include, for example, evaluation information about merchant performance, the information being derived from data processed by a selected processing module.

The foregoing enumeration of embodiments has been for illustrative

10 purposes only. Other embodiments, combinations of embodiments and combinations of features are also within the scope and spirit of the teachings described herein, as will be apparent to persons skilled in the art of these teachings.

15 **BRIEF DESCRIPTION OF THE DRAWINGS**

Figs. 1 (a-g) is an example of a possible consumer survey questionnaire for rating an online merchant and collecting demographic information about a consumer.

Figs. 2 (a-b) is another example of a possible consumer survey questionnaire for rating an online merchant and collecting demographic information about a consumer.

Fig. 3 is a schema of a system for processing and evaluating data relating to e-commerce, in accordance with one or more aspects of the present invention.

Fig. 4 is a block diagram that generally illustrates some features of a computer system that may be used in the present invention.

5 Figs. 5a-e show means of presenting certain kinds of information processed from data collected from an exemplary consumer survey questionnaire, the information being presented in tabular and graphical formats.

Figs. 6a-b show means of presenting certain kinds of information processed from data collected from an exemplary consumer survey questionnaire, the information being presented in tabular and graphical format.

10 Fig. 7 is a flow chart for a processing module that may be used in the system of Fig. 3.

Fig. 8 is a flow chart for steps performed by the processing module of Fig. 7.

15 Fig. 9 is a flow chart of an alarm filter system, according to the present invention, that activates an alarm under specified conditions.

Fig. 10 is a flow chart for a processing module that may be used in the system of Fig. 3.

20 Fig. 11 is a schematic representation of functions related to the processing module of Fig. 10.

Fig. 12 is a graphical representation of functions related to the processing module of Fig. 10.

Fig. 13 graphically illustrates that the prime pull (saturation limit)  $P_0$  draws and meters the growth of  $P_1$ , in accordance with the principles of the processing module of Fig. 10.

Fig. 14 is an input array related to the processing module of Fig. 10.

5 Fig. 15 is a flow chart for a processing module that may be used in the system of Fig. 3.

Fig. 16 is an example web page from an infomediary website showing features of the present invention.

10 Fig. 17 is an example web page from the infomediary website of Fig. 16 showing features of the present invention.

Fig. 18 is an example web page from the infomediary website of Fig. 16 showing features of the present invention.

Fig. 19a-b are an example web page from the infomediary site of Fig. 16 describing certain features in accordance with the present invention.

15 Fig. 20 is an example third party web page that is linked to the infomediary web page of Fig. 18, which is included to illustrate features of the present invention. (The applicant for patent claims no rights in the content of the third party web page shown, with all rights to such content remaining with the owner of that page.)

20 Fig. 20 is an example web page from the infomediary website of Fig. 16 showing features of the present invention.

Fig. 21 is an example web page from the infomediary website of Fig. 16 showing features of the present invention.

Fig. 22 is an example web page from the infomediary website of Fig. 16 showing features of the present invention.

5 Fig. 23 is an example web page from the infomediary website of Fig. 16 showing features of the present invention.

### DETAILED DESCRIPTION OF THE INVENTION

The present invention is a novel system and method for processing and 10 evaluating data, and generating information relating to networked information retrieval and analysis, planning and control; and particularly e-commerce. The system preferably uses data gathered from or about transactions that occur online. The data may be gathered from any number of sources. It may be directly gathered from providers of goods or services, e.g., merchants, and users 15 of goods or services, e.g., consumers, or other participants in an online transaction. The data may also be gathered from third-party suppliers of previously collected data. The data may be evaluated with respect to transactions with particular merchants, a group of merchants, a group of consumers, or with respect to one or more categories of trade based on data 20 gathered from one or more merchants, consumers, or third-party data sources in one or more trade categories.

In addition to evaluation of merchants and any other suppliers of goods and services, and trade categories, the present invention also contemplates evaluations of specific products and services. The principles of the present invention are equally applicable to such evaluations. Products and services may

5 be evaluated based on defined attributes. For example, defined attributes could include: satisfaction with product quality, price, durability, longevity, safety, warranty, customer service; ease of assembly, use, maintenance or repair; cost of operation or repair; upgradability; compatibility, etc. Consumer survey questionnaires served at time of purchase or following a purchase could be used

10 to collect attribute data. In addition, the data could be provided by expert non-consumers such as expert evaluators, as well as other conventional means.

The present invention also contemplates that an intermediary data source could collect data about merchants or products. An intermediary would be a party that is not the manufacturer or seller of the thing being evaluated. For

15 example, an intermediary website could allow evaluators, such as consumers, to complete questionnaires for merchants, products, or services they wish to rate. The questionnaires could be provided on web pages served to consumers by the intermediary's web server. The website could offer incentives to motivate consumers to complete surveys. The intermediary could supply or provide the

20 completed survey questionnaire data to or for the data capture systems of the present invention, as described below.

To illustrate features of the present invention relevant to processing and evaluating data received from the foregoing possible sources, the present invention will be discussed in terms of a specific example. This example is based on online retail transactions between consumers and merchants for goods

5 or services sold online. Persons skilled in the art will recognize that the principles of the present invention illustrated in the example can readily be adapted for other e-commerce applications involving sellers and buyers, including business to business transactions. Accordingly, the embodiments shown in Figs. 1-15, and described herein, should not be construed as limiting

10 the invention to the specific embodiments shown and described.

Figs. 1 (a-g) show one possible embodiment of an online survey questionnaire. In particular, the Figures show screen shots of a consumer survey questionnaire for a retail online transaction. The survey questionnaire may be accessed through a direct link or a framed linked on a merchant's web page that connects to the survey questionnaire host server. The consumer may be invited to fill out an online survey questionnaire following an online purchase, for example. The invitation may be conveniently included on the invoice or order confirmation web page or email.

The survey questionnaire 100 invites a consumer to rate a merchant 20 relative to five attributes in ten dimensions (a 1-10 rating):

(1) *Ease of Ordering* 102;

This attribute relates to the convenience and speed of ordering from the merchant's website.

*(2) Product Selection 104;*

5 This attribute relates to the breadth of products that the merchant has available, keeping in mind the merchant's stated area of focus.

*(3) Product Information 106;*

This attribute relates to the quality, quantity, and relevance of information provided for making a purchase decision an informed one.

*(4) Product Prices 108; and*

10 This attribute considers product prices relative to the competition.

*(5) Web Site Navigation & Looks 110.*

This attribute relates to the overall layout/organization, movement around the site; missing/non-functional links; speed; and how appropriately graphics were used to enhance the shopping experience.

15 The survey questionnaire includes ten merchant rating fields 112 for each attribute. It also includes an "NA" (not applicable) field. A rating of 1-2 corresponds to "not at all" satisfied; 3-4, "a little" satisfied; 5-6 "somewhat" satisfied; 7-8 "quite a bit" satisfied; and 9-10 "highly" satisfied. Performance ratings are intended to give merchants an accurate and comprehensive understanding of how they are performing in the eyes of their consumers.

The collected survey questionnaire data from a survey questionnaire 100 is communicated to a data capture server 14 where the data is parsed, stored, and communicated to storage, processing modules, etc. described herein.

As shown in Figs. 2 (a-b), one or more post-purchase or "fulfillment"

5 survey questionnaires 200 may be sent to a consumer that has filled out the survey questionnaire 100 and taken delivery of the purchased product at set intervals. The survey questionnaire may, for example, be sent by email and may include a hyperlink to a survey questionnaire server 14 where the collected survey questionnaire data is parsed, stored, and communicated to processing

10 modules described herein. The consumer is asked to rate the merchant relative to five attributes in ten dimensions:

(1) *On-Time Delivery* 202;

This attribute relates to timeliness in the context of the promised delivery date.

15 (2) *Product Representation* 204;

This attribute relates to how well the online product description and depiction compared to what was actually delivered.

(3) *Level & Quality of Consumer Support* 206;

This attribute relates to how available and effective the merchant was in 20 resolving any questions/complaints or problems that the consumer encountered.

It also relates to any steps the merchant took to make sure that that the

consumer was informed of order status and was happy with the transaction. (This is left blank if not applicable.)

(4) *Posted Privacy Policies* 208;

This attribute related to the security of credit card, personal, and  
5 transaction information.

(5) *Product Shipping & Handling* 210.

This attribute relates to the appropriateness and condition of product  
packaging.

As with the initial survey questionnaire 100, a merchant's fulfillment ratings

10 may be expressed as a rating on a scale of 1-10 in fulfillment rating fields 212. A  
rating of 1-2 corresponds to "not at all" satisfied; 3-4, "a little" satisfied; 5-6  
"somewhat" satisfied; 7-8 "quite a bit" satisfied; and 9-10 "highly" satisfied.  
Performance ratings are intended to give merchants an accurate and  
comprehensive understanding of how they performed in the eyes of their  
15 consumers.

It is to be understood that survey questionnaires 100 and 200 are merely  
exemplary. The survey questionnaires may include more or less attributes than  
are shown, may be based on other quantitative rating scales, may be based on  
qualitative ratings, etc. In other words, the survey questionnaires may be  
20 customized to evaluate any matter of interest. A survey questionnaire may also  
include fields relating to a consumer's particular demographics, preferences,  
dollar amounts spent, expectations, comments, etc. as shown in survey

questionnaires 100 and 200. Figs. 1b-1f and 2b are screenshots of survey questionnaires that query a consumer on such other information. The survey

questionnaire may include data that is automatically attached to the survey

questionnaire or otherwise associated with it by the merchant. This data may

- 5 include an identifier for the purchased product (e.g., sku), product purchase price, number ordered, etc. The information can provide comprehensive measurement and benchmarking of buyer demographics, transactional behaviors, and shopping preferences because it has been continuously collecting point-of-sale data from online consumers. The foregoing survey questionnaires
- 10 and other potentially obtainable data through such survey questionnaires or other means are examples of data relating to an e-commerce transaction between a consumer and a merchant.

Among other things, the information that may be extracted from the data helps merchants increase buyer conversion and build loyalty through

- 15 understanding buyer preferences; and provides companies with the ability to ask custom survey questions directly from any demographic or online shopping category by appending those questions to a survey questionnaire such as survey questionnaire 100 or 200. The data may also be developed into information that helps educate consumers about their options in the e-commerce marketplace.

- 20 After a consumer inputs data into a survey questionnaire, the captured data is entered into a survey data store and processed into information according to the methods described in more detail below.

As used herein, "data" means facts and beliefs about the real world; "information" is data that has been processed and formatted to support decision making.

Fig. 3 shows an overview of a through-flow system 10, according to the

5 present invention. As used herein, the term "through-flow" generally means a system for gathering, transferring, processing, evaluating data, and/or generating information in a continual stream, as well as presenting and delivering such data and/or information in desired media to interested parties or other systems. The system 10 may be used to evaluate data from a variety of sources, as mentioned  
10 above. As indicated, for instance, it may receive and evaluate data from consumer satisfaction survey questionnaires, it may receive data from online merchants about their activities, or it may receive data from users or third parties who gather and provide data about electronic commerce. Generally, system 10 includes two or more networked computers. System 10 may be divided into a  
15 client or remote system 11 that is composed of one or more remote computer systems or terminals, e.g., 12 and 13, and a server or local system 5 composed of one or more computer systems with one or more sets of system components that are physically located together or networked together over some distance. System 5 is the side of system 10 that collects, processes, and evaluates data, and generates and presents the resulting information. System 11 is the side of  
20 system 10 that may provide data to system 5, or retrieve data from system 5.

Fig. 4 illustrates features that would typically be found in a computer system of

system 10. As used here in a "computer system" generally means a set of components that include one or more of the following: central processing unit ("CPU") 4.1; memory 4.2 and processing modules 22 or user programs 4.21, operating system 4.22 and network interface 4.23, and related I/O subsystems

5 4.3 and 4.4, including one or more of the following: disk drive, keyboard, mouse, display monitor, networking card, other subsystems well-known in the art, and related software applications, including web browsers, web servers, database, and/or communications software. It will be understood by persons skilled in the art, that the computer system, particularly computer systems in system 11 may

10 also be in the form of a Personal Digital Assistant (PDA), a wireless computer system or device capable of network communications over the Internet or other network, or a computer terminal or Internet appliance capable of such network communications.

The local system 5 generally includes data capture server 14 that accepts

15 data transfer from one or more data sources from remote system 11 computer systems or terminals. The data may also enter the data capture server by direct manual input or connection to storage medium such as a CD-ROM, disk drive, floppy drive, memory cards, ZIP drive cartridges, and the like. System 5 further includes database 16, data conditioner 18, database 20 for conditioned data, one

20 or more data processing modules 22, a display store 24 for storing processed data, and a presentation server 26 for presenting data or information to an end-user. These, and other possible software components of through-flow system 10

described herein, reside on one or more hardware components providing addressable memory for a computer system of local system 5.

Computer systems of system 10 may execute a conventional operating system, including, but not limited to, Microsoft Corporation's Window (sm) 3.x,

5 95, 98, 2000, NT or DOS operating systems; Linux operating systems; Sun Microsystems' Unix Solaris operating systems; IBM's AIX Solaris operating systems; SGI's IRIX Solaris operating systems or other operating systems, including, but not limited to Window's CE, Palm OS, etc. for PDAs, cellular devices, web devices, or other computing devices.

10 The computer systems on system 10 may access and store data in any variety of data storage media, including, but not limited to, local hard drive, CD-ROM or other mass storage device, local area network servers, wide area network servers, Internet servers accessed via phone or cable, commercial consumer and business online services, distributed information sources such as

15 any combination of the above, and local RAM if data is generated on demand or linked to other programs such as spreadsheet or database programs executing in RAM.

The components of computer systems of remote system 11 of the present invention may be disposed at various points of use including, but are not limited to, the home (such as on a stand-alone multi-media PC connected to an online service), office (such as on a workstation computer connected to a database on a WAN), kiosk used in-store, in hospitality services or training environments.

Turning to the specifics of system 10, the data that enters data capture 14 may relate to an online commercial transaction such as the type discussed above relative to survey questionnaires 100 and 200. Or it may be any other data type sought to be collected. For example, a consumer making an online purchase of

5 goods from a merchant's website may be invited to fill out a consumer satisfaction survey questionnaire 100 or 200 following the transaction. The merchant's website may be located on a web server which may be part of or separate from system 10. In the case of online survey questionnaires, the survey questionnaire may be provided to the consumer as HTML, XML, Java

10 Script, ActiveX, Applet, or other well-known mechanisms or formats for querying a local computer system through a web browser. (Other servers in system 10 may use the same mechanisms or formats for interfacing with other computers.) The survey questionnaire could be served by a web server that is part of local system 5. For example, the web server could reside on data capture server 14

15 or presentation server 26.

In local system 5, data capture server 14 serves as a web server for capturing the data that a consumer inputs into consumer satisfaction survey questionnaire 100. In general, a web server is a computer system that runs software that enables access by remote computer systems or terminals via the

20 Internet or other network transport to allow presentation of information, data storage and transfer, and other transactions. Suitable web server software is well known and includes Apache Software Foundation's Apache Server web

server; Netscape's FASTTRACK series web server, Microsoft Corporation's IIS, and Sun Microsystems' JAVA WEB SERVER.

The survey questionnaire may be presented to a consumer contemporaneously with a transaction or at some defined time after the

5 transaction, or both. Email is another contemplated mode of communicating a survey questionnaire to a consumer. The email could contain a plain text based questionnaire that would be parsable by data capture server 14. Or the email might contain a link to an address that could serve the survey questionnaire in HTML, XML, Java Script, ActiveX, Applet, or other well-known mechanisms or

10 formats for querying a consumer through a web browser.

In one possible embodiment, the invitation to fill out survey questionnaire 100 is found on the receipt page that a consumer receives confirming an online order. The invitation includes a hyperlink to a web server, e.g., data capture server 14, that presents survey questionnaire 100 via the browser of a consumer

15 computer system 12.

It is also possible that the data collection process could be interactive. For example, the data received for certain fields may prompt a web server or other component of system 10 to query the consumer, or other provider of data, with respect to new fields that are dependent on data values sent to the system

20 server 14. A survey taker might be asked to identify themselves as a male or female, with further survey questions being selected based on the answer.

The survey questionnaire could contain defined fields for data values. For example, survey questionnaire 100 could be designed to accept data values that represent a rating on a 1-10 scale of the consumer's perceptions about defined aspects of a transaction with a merchant, as described above relative to survey 5 questionnaires 100 and 200.

In addition to consumer perceptions about aspects of a transaction, the survey questionnaires could seek any other kind of information in which a survey administrator may be interested. Such information may include demographic information about a consumer (e.g., age, profession, education, etc.); the

10 consumer's perceptions about, and experiences with, online shopping (not limited to the transaction conducted); the consumer's interest in other lines of goods and services; whether the consumer wishes to participate in newsgroups, future offerings, etc. relating to the purchased product or any other subject matter.

The means for obtaining and collating survey data is well known in the art.

15 For example, such techniques are disclosed in Peters, et al., US Patent No. 5,893,098 entitled "System And Method For Obtaining And Collating Survey Information From A Plurality Of Computer Users", the entire contents of which are hereby expressly incorporated by reference. For purposes of illustration, the following description shall be in terms of an online merchant rating survey 20 questionnaire 100, as described above. The data shall be discussed in terms of a 1-10 rating entered relative to attributes 102-110 in rating fields 112.

After a consumer completes a survey questionnaire, the resulting survey data is fed to data capture server14 and logged into a server log. The log records the collected data and may assign an identifier or key value to the data.

For example, the identifier could be an indicator of any number of things

5 including the time of receipt, source (e.g., the particular merchant), etc. The logged data may then be parsed by a software program that assigns items of the raw data into predefined fields in a database 16. The parsing program may be created through standard program techniques that are known or within the skill of persons in the art. The parsing program would be designed to map data from the

10 survey questionnaire to database records that are predefined.

Each record would be composed of one or more fields that correspond to the class of an item of data and/or the value of an item of data. Any field may be keyed to any one or more other fields. For example, in the case of survey questionnaire 100, the database would contain tables that include a field for each 15 particular attribute102-110. The numerical rating from each survey questionnaire, or average rating from a predefined group of survey questionnaires, would be listed as items under the relevant attribute 202-210 fields. Each table would typically include other fields such as survey identifier, time of receipt, source, etc.

20 Database 16, as well as any other database contemplated by this invention, including databases 20 and 24, may be based on any number of known database management systems (DBMS), including hierarchical

databases, network databases, relational databases, and object oriented databases. Suitable DBMS are widely available and include Oracle, Sybase, Microsoft SQL Server, and DB2 . One suitable database system is a relational database based on SQL language. A suitable SQL database would typically

5 include a Data Definition Language (DDL) for creation of a new database and new objects within an existing database; a Data Manipulation Language (DML) for processing existing database objects; and a Data Control Language (DCL) for administering permissions and other security-related issues. The data processing is handled by a database engine database and can be accessed by various

10 searching means, including Boolean logic evaluation, proximity calculations, and fuzzy logic evaluation. The databases of the present invention may also be custom designed, the design process and features being well within the skill of persons in the art.

The data fielded into database 16 may be conditioned by a data

15 conditioner 18. The data conditioner 18 is one or more applications that perform some operation relative to the data. Data may be conditioned to clean, filter, or otherwise to process data. For example, the data may be conditioned to supply imputed and/or default values to data items missing from a particular field or to eliminate anomalous data that does not meet predefined criteria in terms of class

20 and/or value. In further example, the data may be conditioned to supply missing values; to eliminate data collected during holidays or other non-business days; to eliminate data based on demographic criteria; to check for anomalies; etc.

Standard software programming skills may be used to develop desired data conditioning programs.

Conditioned data may be stored in a database 20 for conditioned data.

The data collected by the database 16, or other contemplated databases, may be

- 5   communicated to the data conditioner 18, or other system components such as database 20 or processing module(s) 22 (discussed below), in a continuous stream, in batches at set intervals, or in batches. Communication of data could be made dependent on specified conditions. For example, the specified conditions could include a set quantity of data having been received (e.g., after
- 10   receipt of one hundred consumer responses to survey questionnaires); an elapsed period of time (e.g., after one hour of collecting responses to survey questionnaires), or other specified conditions or combinations of such conditions.

Databases 16, 20, and 24 are not intended to be limiting examples of databases that are used in the present invention. Additional databases may be

- 15   incorporated into system 10 besides those specifically indicated. These additional databases may be for the receipt of raw data, conditioned data, or processed data that flows through system 10, or selected categories of such data. It should also be appreciated that the databases may reside on the same hardware components or different hardware components. The databases could
- 20   also be operated under the same or different DBMS.

One or more data channels 15 may be defined for the transmission of data between the components of system 10. As used herein "data channel" means a

logical event, not necessarily a physical channel, directing the flow of data into computer storage and/or processing systems. Data channels are determined by the structure of, for example, a survey questionnaire, or the way answers are coded into database fields, and the programs that are to operate on predefined 5 categories or items of data. A data channel may also be a logical or computed combination of other data channels.

Looking at Fig. 3, one or more data channels 15 communicate parsed data from server 14 to assigned fields in database 16. More specifically, there could be a data channel assigned to each attribute 102-110 for communicating 10 the ratings from rating fields 112 into the field in the database that correspond to each attribute 102-110. Similarly, one or more data channels 15 may communicate data from defined fields in database 16 to data conditioner 18. For example, raw data from database 16 or conditioned data from database 20, may be introduced into database 20 or directly into a processing module(s) 22 15 consisting of one or more processing modules that perform a predetermined operation on data. The processing modules may be in communication with each other or with other system components. As shown in Fig. 3, processing module(s) 22 transfer processed data to a database 24 for storing processed data. Database 24 is in communication with a server 26. Server 26 may be a 20 web server for interfacing with remote computer systems in system 11. Web server 26 may be used to present data from database 24 to remote computer systems 12 or 13 over a network 11. Alternatively, presentation server 26 could

present the data to a printer 34 or video display or computer system 36, or through other means for displaying data. (As used herein, "data" includes data in the form of files or data representing information, unless otherwise noted or clear from the context.)

5 One or more data channels 15 may communicate data from database 20 to processing module(s) 22. Although not illustrated in all cases, it will be understood that any other components of system 10 may communicate defined fields of data to another system component along a data channel assigned to a field(s) of data.

10 It is to be understood that the data channels depicted in Fig. 3 are for illustrative purposes only and should not be construed as a limitation on how data may be communicated between system components, other paths being within the spirit and scope of this disclosure. For example, data conditioner 18 is optional; although not illustrated, data may be directly introduced from database 15 16 into a processing module 22.

Processing modules 22 may process data communicated to it to generate desired forms of information. The processing may include computation of ratings; indices; consumer demographic profiles; merchant or merchant category session volumes; dollar volumes; lists of consumers for target marketing; etc. As 20 used herein, such data or information may be referred to as "e-commerce" data or information. Standard software for such processing is well known and

available, or may be readily created by persons skilled in the art, and therefore is not discussed further herein.

In the case of survey questionnaire 100, the processing module 22 could be programmed to compute ratings for each attribute 102-110, and other

5 standard statistical values based on the ratings data. Accordingly, a statistical analysis process module would be capable of computing one or more of the following:

- Means using real numbers
- Mean scores using factors (useful for rating scales)
- Standard deviation
- Standard error
- Error variance
- Z-tests (four tests for the difference between proportions)
- T-tests (two tests for the difference between means)
- P-values (a test of the probability that an event was chance)
- Significance net difference test (on pairs of columns)
- Least significant difference test (on means)
- Paired preference test (on pairs of rows)
- Chi-squared tests (one or two dimensional and single classification types)
- Proportions tests (four types)
- Friedman's two-way analysis of variance
- Kolmogorov-Smirnov test (on differences between two samples)
- McNemar's test of the significance of changes
- F-test for testing differences between a set of means
- One-way analysis of variance (ANOVA)
- Significance levels
- Regression analysis
- Trend analysis
- Correlations
- Covariances

The software for statistical analysis could be designed using programming skills well within the skills of persons in the art. Well-known, off-the-shelf programs are also available. Such programs include the SPSS family of software products, including SPSS BASE PROFESSIONAL, v10.5 or v8.0, and

5 Regression, Table, and/or AnswerType module from SPSS, Inc., Chicago, IL.

The processed data may then be transferred to a "display store" database 24 for storing processed data or information. Database 24 may be connected to a presentation server 26. The presentation server 26 presents data that has

been processed and evaluated by system 10 to an end-user, such as consumers

10 or merchants 13. The display store information may be accessible publicly or privately via computer systems in system 11. Server 26 parses and sorts

processed information into categories that are defined according to defined needs or interests of an end-user or group of end-users. The software for doing

this must be coded according to the intended application of the user. The coding

15 may be written in standard programming codes that are known to persons skilled in the art. The relevant information may be presented to the end-user in any

known form of communicating information, including: printed reports, emailed

reports; server-stored reports accessible to end-users over a public or private

network, including website posted reports. The presentation server may serve

20 web pages to remote computers systems through conventional TCP/IP protocols or other known protocols. The web pages may be presented through conventional web browsers such as INTERNET EXPLORER series of web

browsers by Microsoft Corporation, or the NETSCAPE COMMUNICATOR series of web browsers by Netscape Communications Corporation.

The following section describes an embodiment of system 5 that is adapted to process data from survey questionnaires 100 or 200 from a sample of 5 survey questionnaires from consumers that did business with online merchants.

The survey questionnaires may be supplied to data capture server 14 at predefined intervals, in real time, or upon other specified condition. As will be clear after the detailed discussion of modules 221-223 provided below, the present invention can allow daily or more frequent updates of usable information

10 to end-users even though only small and/or noisy data samples are available.

Accordingly, the present invention sets new standards for minimizing the "time late" for turning data into information.

Figs. 5a-e show examples of information that may be derived from the ratings from a set of survey questionnaires 100/200 for a particular merchant for 15 transactions occurring over set intervals (e.g., one day, one week, one month, etc.). Referring to Fig. 5a, a table contains the average numerical rating for each attribute 102-110/202-210 on a monthly basis over one quarter. Figs. 5b and 5c show in graphical form the average ratings from Fig. 5a. The graphs enable users to easily detect upward or downward trends for each attribute. Like Fig. 5a,

20 Fig. 5d shows a table with the average quarterly rating for each attribute shown in Fig. 5a. Fig. 5e is a graphical representation of the quarterly average ratings for each attribute from Fig. 5d.

For simplicity, the following discussion shall be in terms of the *Product* Price attribute 108, although it should be understood that the principles discussed are applicable to any other attribute 102-110/202-210.

Turning now to Figs. 6a-b, specific information about price attribute 108

5 from Figs. 5a-e is shown in table and graphical formats. The table shows five categories of ratings. A rating of 1-2 corresponds to "very low satisfaction"; 3-4, "low satisfaction"; 5-6 "moderate satisfaction" 7-8 "high satisfaction"; and 9-10 "very high satisfaction".

10 DATA STABILIZER

Figs. 7-8 and equations 1-34 below relate to a novel application for estimating or smoothing variable data. The schema presented below is particularly useful for estimating values based on noisy and/or small data samples. It may also be used to predict missing observations. The schema may 15 be incorporated into system 10 as a processing module 22. Through the schema, the present invention offers notable advantages over traditional modes of statistical analysis, such as moving window averages (weighted and unweighted), which do not provide accurate information if data samples are small or noisy. Consequently, in conventional techniques, the conversion of data to 20 useful information is prolonged by the time necessary to collect the larger required sample sizes. Thus, the schema of the present invention provides notable advantages over conventional methods used to evaluate data related to

e-commerce. In one possible embodiment, the schema is used to stabilize or filter consumer ratings data. While, for convenience, the following discussion is largely in terms of evaluation of ratings data, the present invention also contemplates use of the schema in evaluating other forms of data related to e-commerce. Accordingly, persons skilled in the art will appreciate that the present invention provides a substantial and patentable advance over prior methods of processing and evaluating such data, particularly in the area of business ratings, market research studies, consumer demographics, etc.

*S<sub>1</sub> B<sub>1</sub>* A central component of the schema of Fig. 7 is an Extended Kalman Filter or ("EKF"). The EKF is detailed in Fig. 8. Generally, the EKF uses a computational (recursive) solution of the minimum variance Bayesian estimation method. The EKF is powerful in several aspects: it supports estimations of past, present, and even future states. It can do so even when the precise nature of the modeled system is unknown. In addition to smoothing noisy data, evaluating small samples of data, and providing a basis for estimations, the EKF also provides a method of weighting data values according to the recency or level of noise corruption of the data. This may be important because, for example, data collected in a later portion of a data collection interval (data window) is likely to be more indicative of present trends than is data collected at the earlier portion of the interval (older data).

Continuing with the example of survey questionnaire 100, the present invention provides a sequential filter that outputs an optimal estimate of a "rating"

for an attribute 102-110 from a set of responses to survey questionnaires of such rating from a given merchant. The following section describes how the true rating may be determined from noisy data/small sample sizes. (In this section, related Figures, and in subsequent sections, certain mathematical equations are

5 identified by a number in bold and in brackets to the right of a given equation, which number should not be confused with unbracketed reference numerals for the accompanying Figures.)

~~S62~~ In one preferred embodiment, a true rating  $A$  is determined using an adaptation of an Extended Discrete Kalman filter. It is to be understood that the

10 following embodiment is presented for purposes of illustration not limitation. Persons skilled in the art will appreciate that other adaptations of Kalman filters are within the scope and spirit of the present invention. In connection with the following discussion, reference may be made to Figs. 7 and 8, which help illustrate the principles being discussed. Hereinafter, a processing module that

15 can smooth noisy or variable data using a computational (recursive) solution of the minimum variance Bayesian estimation method is referred to as a Data ~~Stabilizer or "DS" for short.~~

It may be assumed that in a noisy sample of ratings observed over a set time interval, the desired "true rating" is actually a slowly varying function of time, 20 while the noisy (input) rating suffers from "sampling jitter" induced by often low response (small sample) and/or highly variable ratings for the input parameters from which the rating is computed.

The true rating  $A$  is a scalar valued function of parameter vector  $y$  of dimension  $N$ .

$$A = f(\underline{y}) \quad [1]$$

where  $\underline{y}$  is a vector of sampled variables with mean  $\hat{\underline{y}}$  and possibly correlated

5 dispersions covariance  $\underline{y}$ . Once  $A$  is determined, it could be reported as one of the attribute ratings shown in Figs. 5-6 or used in further computational processes of system 10 or elsewhere. To determine  $A$ , it may be assumed that variations of  $A$  over time can be described by a second or higher order polynomial in time  $t$ .

$$10 \quad A(t) = c_0 + c_1 t + c_2 t^2 \quad [2]$$

$$= \underline{x}^T [t^0, t^1, t^2]^T \quad [2.1]$$

$$\underline{x} = [c_0, c_1, c_2]^T \quad [3]$$

$c_i$  are the components of  $\underline{x}$ , the state vector of constants for the polynomial [2].

A higher order polynomial for  $A(t)$  can be used if warranted by the underlying process modeled. (T means "matrix transpose", as is known in the art.)

**Observation Process:** This "ratings process" is actually observed by a

15 sequence of computed noisy ratings 221.2 at times  $t_i$ ,  $i = 1, 2, 3 \dots$  such that

$t_{i+1} > t_i$ . Then

$$\tilde{A}(t_i) = h[\underline{x}(t_i)] \quad [4]$$

with

$$\text{cov } \tilde{A}(t_i) = \left[ \frac{\partial h^T}{\partial y} \text{cov } y \frac{\partial h}{\partial y} \right]_{\tilde{y}(t_i)} \quad [5]$$

$\tilde{A}$  is termed the pseudo observable; it is a function of actual observations. In what follows, the time index  $i$  may be used to denote the value of a time indexed variable at  $t_i$ .  $\tilde{y}$  is the observed parameter vector which itself may be computed

5 from captured survey questionnaire responses.

**State Vector:** The estimation problem may now be reduced to finding the optimal value  $\underline{x}$  of  $x$  which is termed the state vector of the process. Then at  $t_i$  we have

$$\underline{x}_i = [c_{0,i}, c_{1,i}, c_{2,i}]^T \quad [6]$$

10 with covariance  $P_i = \text{cov } \underline{x}_i$  [7]. It is noted that  $t_{i+1} = t_i + \Delta t_i$  for all  $i$  and arbitrary  $\Delta t_i > 0$ .

**State Transition:** The estimated polynomial [2] is assumed to be stable over "reasonable" time intervals that encompass the last few observation times. Accordingly, the state transition equation may be written as

15  $\underline{x}_{i+1} = I \underline{x}_i + \underline{r}_i \quad [8]$

where  $I$  is the identity matrix and  $\underline{r}_i$  is "state noise" such that  $E(\underline{r}_i) = 0$  and  $R_i = \text{cov } \underline{r}_i$  [9]. It may be further assumed that  $\Delta t_i$  are large enough so that  $E(\underline{r}_i^T \underline{r}_j) = 0$  for  $i \neq j$ , i.e. that the noise is uncorrelated over time (explicitly  $E(\underline{r}_i \underline{r}_j^T) = R_i \delta_{ij}$ )).

**Observation Equation:** The pseudo observation  $\tilde{A}_i$  then is related to state by

$$\tilde{A}_i = \underline{m}_i^T \underline{x}_i + w_i \quad [10]$$

where the pseudo-observation or measurement vector  $\underline{m}_i$  is given by

5.  $\underline{m}_i = [1, t_b, t_i^2]^T \quad [11]$

and the measurement noise is

$$E(w_i) = 0 \text{ and } E(w_i w_j) = W_i \delta_{ij} \text{ for all } i, j. \quad [12]$$

It is noted that the observation covariance  $W_i$  is identically the error induced in  $\tilde{A}$  by the errors propagating from the actual observed parameters  $\tilde{y}$  given in [5].

10 Therefore we set

$$W_i = \text{cov } \tilde{A}_i = \left[ \frac{\partial h^T}{\partial y} \text{cov } y \frac{\partial h}{\partial y} \right]_{\tilde{y}_i} \quad [13]$$

The estimator elaborated below is the discrete Kalman filter, as described for example, in 1 Brogan, W., *Modern Control Theory*, Prentice-Hall, 1985. The discrete Kalman filter is here adapted to the observed ratings process developed above. In the following description, the subscript  $j/i$  denotes a parameter value at  $t_j$  computed by input(s) available at  $t_i$ . The filter gain matrix (in this case a vector) for processing the input at  $t_i$  is written as

$$K_i = P_{i|i-1} \underline{m}_i [\underline{m}_i^T P_{i|i-1} \underline{m}_i + W_i]^{-1} \quad [14]$$

where the state noise propagates as captured by the covariance matrix  
20  $P = \text{cov}(\underline{x})$  over  $\Delta t_{i-1}$  as

$$P_{i|i-1} = P_{i-1|i-1} + R_{i|i-1} \quad [15]$$

If we let  $C_i = I - K_i \underline{m}_i^T$  [16] then the covariance of state at  $t_i$  (i.e.  $\text{cov } \underline{x}_{i|i}$ ) is given by

$$P_{i|i} = C_i P_{i|i-1} C_i^T - K_i W_i K_i^T \quad [17]$$

5 With these preliminaries, we write the estimate of state for  $t_i$  now as

$$\hat{\underline{x}}_i = \hat{\underline{x}}_{i-1} + K_i [\tilde{A}_i - A_{i|i-1}] \quad [18]$$

(this represents the main equation for Kalman filtering),

noting that  $A_{i|i-1} = A_i(\hat{\underline{x}}_{i-1}) = \underline{m}_i^T \hat{\underline{x}}_{i-1}$  [18.1] is the predicted rating. (We use the single subscripts in [18] and elsewhere for simplicity due to the constancy of  
10 state evolution as indicated in [8] above.)

From [15] we see that  $R$  forms a floor for the error covariance of our estimate  $\hat{\underline{x}}$  since at every stage we propagate error by adding  $R$  to the last estimated covariance. We also see that the filter gain  $K$  is diminished by encountering a very noisy observation represented by a large  $W$ . This lets  $R$  be  
15 used as the filter memory length controller, i.e. we should believe time late estimates only to the extent that we believe in the validity (or accuracy) of our state transition model [8]. This lets us shorten the memory length by increasing  $R$  thereby appropriately abandoning or deweighting past observations.

We recall that the estimate is updated in [18] by multiplying the currently  
20 observed "innovation" ( $\tilde{A}_i - A_{i|i-1}$ ) by the filter gain, where "innovation" is the difference between what we actually observed and what we expected to observe.

If our state model is good and our filter is working properly, then the innovation sequence will approach a zero-mean white noise process with covariance given by [5].

Finally, the data stabilizer and its variance is computed from

5  $\hat{A}_i = \underline{m}_i^T \hat{x}_i$  [19]

and  $\sigma_i^2(\hat{A}) = \left\{ \left[ \frac{\partial A}{\partial \underline{x}} \right]_i^T P_{i|i} \left[ \frac{\partial A}{\partial \underline{x}} \right]_i \right\}$  [19]

$= \underline{m}_i^T P_{i|i} \underline{m}_i$  [20]

Implementation: The time origin may be set to always represent the 10 present time, i.e.  $t_i=0$ . This requires some important modifications to the above-derived model, which will make the resulting computations more efficient.

In [8] the transition will now require the use of a non-trivial transfer function

$\Phi(i, i+1)$ . This is derived as follows:

15 Let  $t_{i-1}=0$  (i.e. at last estimate of polynomial), then using  $\underline{x}_{i-1}$  we can predict  $A$  ( $t > t_{i-1}$ ) as

$A(t) = [1, t, t^2] \underline{x}_{i-1}$  [21]

Now if we shift the time origin by  $\Delta t_{i-1}$  to  $t_i$ , then the new coefficient vector  $\underline{x}_i$  must satisfy

20  $A(t) = [1, t - \Delta t_{i-1}, (t - \Delta t_{i-1})^2] \underline{x}_i$  [22]

Equating the corresponding time shifted coefficients then gives

$$\left. \begin{array}{l} c_{i-1,0} = c_{i,0} - c_{i,1}\Delta t_{i-1} + c_{i,2}\Delta t_{i-1}^2 \\ c_{i-1,1} = c_{i,1} - 2c_{i,2}\Delta t_{i-1} \\ c_{i-1,2} = c_{i,2} \end{array} \right\} \quad [24]$$

which may be written as

$$\underline{x}_{i-1} = \Phi(i, i-1)\underline{x}_i \quad [25]$$

5 where  $\Phi(i, i-1)$  is now the system transfer function retrodicting state from  $t_i$  backwards to  $t_{i-1}$ .

$$\Phi(i, i-1) = \begin{bmatrix} 1 & -\Delta t_{i-1} & \Delta t_{i-1}^2 \\ 0 & 1 & -2\Delta t_{i-1} \\ 0 & 0 & 1 \end{bmatrix} \quad [26]$$

Invertible system transfer functions have the property that  $\Phi(i-1, i) = \Phi^{-1}(i, i-1)$   
10 [Brogan, p. 291] which we require since we want to go from  $\underline{x}_{i-1}$  to  $\underline{x}_i$ . Therefore,

$$\Phi_i = \Phi(i-1, i) = \begin{bmatrix} 1 & \Delta t_{i-1} & \Delta t_{i-1}^2 \\ 0 & 1 & 2\Delta t_{i-1} \\ 0 & 0 & 1 \end{bmatrix} \quad [27]$$

which gives us the desired state transition and lets us rewrite [8] as

$$\underline{x}_{i+1/i} = \Phi_i \underline{x}_{i/i} + \underline{r}_i \quad [28]$$

The pseudo-observation vector in [11] now becomes simply

$$\underline{m} = [1 \ 0 \ 0]^T \quad [29]$$

since we are "observing" only  $c_{i,o}$  at  $t_i = 0$ .

The estimator equations must now be expanded to include  $\Phi_i$ . Specifically

5 [15] now becomes

$$P_{i|i-1} = \Phi_i P_{i-1|i-1} \Phi_i^T + R_i \quad [30]$$

which changes filter gain [14] to

$$K_i = [\Phi_i P_{i-1|i-1} \Phi_i^T + R_i] \underline{m} [\underline{m}^T (\Phi_i P_{i-1|i-1} \Phi_i^T + R_i) \underline{m} + W_i]^{-1} \quad [31]$$

10 And since  $t_i = 0$  at all estimation points, we have the output of data stabilizer [19]

as

$$\hat{A}_i = \hat{x}_{1,i} = \hat{c}_{0,i} \quad [32]$$

with variance from [20] reducing to

$$\sigma_i^2(\hat{A}) = (P_{11})_{ii} \quad [33]$$

15 where  $\hat{A}_i$  is obtained from the current estimate of state now given by

$$\begin{aligned} \hat{x}_{ii} &= \hat{x}_{i|i-1} + K_i [\tilde{A}_i - A_{i|i-1}] \\ &= \hat{x}_{ii-1} + K_i [\tilde{A}_i - (\hat{c}_0)_{ii-1}] \quad [34] \end{aligned}$$

Fig. 7 shows an implementation of the DS into system 5 as a processing module 221. Observed ratings from a survey questionnaire 100 or 200 at a time  $t$  are stored in database 20, as described above relative to Fig. 3. A nominated response set is defined in step 221.6 of processing module 221, which in Fig. 3 is

represented as one of the processing modules 22. The nominated response set, for example, may be the set of individual ratings for a particular survey questionnaire attribute collected from a particular merchant over a defined interval, such as 24 hours. The nominated response set may be expressed as  $Y_i$

5      $= \{y_j: j=[1, N]\}_i$ . A control parameter database 221.16 is used to store process parameters for process module 221. Step 221.8 calls on process 221 to make a decision relative to  $N$  for the nominated response set. For  $N$  greater than 0, the average value and variance for the set is computed in step 221.10. The effective observed covariance  $W_i$  for the values determined in step 221.10 are then

10    determined in step 221.14.  $W_i$  is then introduced into the EKF steps 221.22, described above, and shown in flow-chart form in Fig. 8. Returning to decision step 221.8, for  $N$  equal to 0, step 221.12 extrapolates the last state estimate by setting  $K_i = 0$ , which outputs to EKF steps 221.22. EKF steps 221.22 then provide outputs to 221.24 wherein the form of the stabilized or true rating  $A$

15    (which may also be expressed as Variable  $v_{\text{stabilized}} = x_{i/I}(I)$ ) and a standard deviation of the data stabilizer (which may be expressed as the square root of  $[P_{i/I}(I, I)]$ ) is extracted. A data stabilizer may then be introduced into a stabilized values database 221.6. This database may be the same as or different from database 24 in Fig. 3. The data stabilizer  $A$  or other output of steps 221.22 and

20    221.24 also may be introduced into other computational processes, e.g., other process modules 22 and/or presentation means, such as presentation server 26.

Data stabilizers may be introduced into control parameters database 221.16 to update it. Averages and variances from process step 221.10, and observed covariances from process step 221.14 may also be introduced into database 221.16 to update. Using control parameters, namely  $R$  &  $W$  matrices,

5 and specified innovation thresholds from database 221.16, an innovation analysis may be performed in step 221.18 on the  $\{\tilde{A}_i = (\hat{c}_0)_{i/i-1}\}$  sequence. Output from step 221.18 may then be used to compute and reset the memory length control  $R_i$  (the covariance of the state transfer function) for input to the EKF steps 221.22. Extended Kalman Filter steps 221.22 may also directly

10 access control parameters database 221.16 to obtain inputs to  $t_b$ ,  $t_{i-1}$ , state vector  $x_{i-1/i-1}$  and covariance  $P_{i-1/i-1}$  for computing the next state vector  $x_{i/i}$  and the covariance of state  $P_{i/i}$ .

Referring to Fig. 8, the general EKF steps 221.22 are shown in more detail. Steps 221.23-221.36 represent an observed process having a state  $x$ . The

15 state of a process may be denoted as  $x_{i+1} = f(x_{i/b}, t_{i-1}, t_i) + r_i$ , where  $t_i$  represents a point in time. In the real world, the estimate of state  $x$  is imperfectly known. In step 221.24  $y^*$  represents an average of a set of one or more noisy observations and  $w_i$  is the observed covariance. The symbology used in Fig. 8 generally corresponds to that of Equations 1-34. (EKF 221.22 is adapted for use in

20 determining true rating  $A$  in Equations 1-34 above, accordingly the Equations 1-34 may not directly correspond to Fig. 8 in terms of symbology, but any

differences in light of the included legend will be understood by persons skilled in the art).

The foregoing estimation model may be implemented in standard programming languages for mathematical functions. Below is one example of 5 the DS model coded in MATLAB programming language. Persons skilled in the art will appreciate that other programming languages may also be used to implement the foregoing principles. The program is used to stabilize ratings from a consumer survey questionnaire:

```
10 clear all
11
12 % read and parse data array
13
14 load daily_aval_d      % daily_aval_d = (value, variance, julian date)
15 nmax = size(daily_aval_d,1); % number of data points
16
17 At = daily_aval_d(:,1);
18 W = daily_aval_d(:,2);
19 wmin = 6; % minimum allowed observation noise
20 wmax = 30; % maximum allowed observation noise
21 W(W>wmax) = wmax;
22 W(W<wmin) = wmin;
23 t = daily_aval_d(:,4);
24
25 % initialize filter
26 %xlast = [At(1) 0]'; % initial state (quadratic coeffs)
27 xlast = [6.5 0 0]';
28 %Plast = [1.5 0 0
29 %           0 .01 0
30 %           0 0 .001]; % covar of initial state
31 Plast = [0.795 0.0082 0
32           0.0082 0.0001 0
33           0 0 0];
34 m = [1 0 0]; % measurement vector
35
36 xsigs = [1e-4 1e-5 1e-4]; % coeff DAILY transition sigmas
37
38 R = diag(xsigs.*xsigs,0); % constant DAILY state transition covariance
39 I = eye(3);
40
41 % Output array initialization
42 StabilizedRatingsout = zeros(nmax,4);
43 StabilizedRatingsout(1,1) = xlast(1); % initialized input value
44 StabilizedRatingsout(1,2) = sqrt(Plast(1,1)); % sigma of input value
45 StabilizedRatingsout(1,3) = t(1); % time value
46
47 % run main loop and generate output array
48 for i = 2:nmax
```

```
dt = t(i)-t(i-1); % current time increment (should be 1)
U = [1 dt dt^2
      0 1 2*dt
      0 0 1]; % state transition matrix
5
if At(i)==99 % then extrapolate from last estimated state
  StabilizedRatingsout(i,1) = [1 dt dt*dt]*xlast; % extrapolated estimate of
value
10  Pnow = U*Plast*U' + R; % extrapolated state covariance
  xnow = U*xlast; % extrapolated state
  else % run filter with current observation
15  w = W(i); % pseudo-observation variance
    Pnext = U*Plast*U' + R; % predicted state covariance
20  temp1 = Pnext*m';
    K = temp1/(m*temp1 + w); % filter gain
    C = I - K*m;
    Pnow = C*Pnext; % current cov of estimated state
25  xnext = U*xlast; % predicted state
    innov = At(i) - xnext(1); % innovation
    xnow = xnext + K*innov; % current estimate of state
30  StabilizedRatingsout(i,4) = innov;
  end % conditional extrapolate or filter calcs
35  % hard limit c0 to max allowed value of rating
  if xnow(1)> 10
    xnow(1)= 10;
  end
40  xlast = xnow;
  Plast = Pnow;
  % fill output array
  StabilizedRatingsout(i,1) = xnow(1);
  StabilizedRatingsout(i,2) = sqrt(Pnow(1,1));
45
end % main observation/time loop
50 % compute 25% level At Risk Threshold ART array
tART = 6.5; % the selected rating threshold
pART = 0.9995; % the probability that actual rating < tART
limART = norminv(pART,0,1);
rART = StabilizedRatingsout(:,1) + limART*StabilizedRatingsout(:,2);
ART = tART*(rART<tART); % array of threshold violations
55 % output results plot.
At(At>10) = 0; W(W>10)=0.5;
  figure(1)
  clf
60  hold on
  tstart = 245; % full range 245 to 638
  tend = 638;
  istart = find(t==tstart); iend = find(t==tend);
  indexART = find(ART>0 & t>=tstart & t<=tend);
65
```

GANZLAW, PC  
PO Box 10105  
Portland, Oregon 97296  
Phone: (503) 224-2713  
Docket No.: BIZ/99-0008  
Express Mail No: EL326146496US Deposited March 29, 2000

```
plot(t(istart:iend),At(istart:iend),'gx',
5 t(istart:iend),StabilizedRatingsout((istart:iend),1),'b',
t(istart:iend),StabilizedRatingsout((istart:iend),2),'b--')
plot(t(indexART),ART(indexART), 'r*',
t(istart:iend),0.1*StabilizedRatingsout((istart:iend),4),'c.' )
grid
hold off
```

10 Copyright in the foregoing code is claimed by BIZRATE.COM.

## 15 ALARM FILTERS

Along any data channel, there is a normative rate of data transmitted and normative values for the data. One or more alarm filter(s) 28 can be set to monitor data rates and send a signal based on deviations from desired thresholds from the normative rate. The normative rate may be based on any 20 defined interval (the "synoptic response rate"). An alarm filter may check for deviations calculated from a sampling of data taken within the normative interval (the "topical response rate"). To set an alarm, a historical or expected arrival rate is determined or set for one or more data channels. The standard deviation may also be determined or set for the arrival rate. A predetermined threshold may be 25 set relative to a topical response rate. The threshold may be set by a system administrator or automatically determined through a selected correlation function. If the alarm filter determines that a threshold is crossed, it triggers an alarm that may be sent to an event handler.

For example, a normative interval may be set at one month as the period 30 over which a consumer survey questionnaire is collected from a particular

merchant's website. Suppose the response rate for a particular merchant's consumer survey questionnaire averages 300 per 30-day period. This average is the historical arrival rate. This gives an average of 30 surveys collected per day.

An alarm filter could be set to monitor the number of surveys collected per day

5 from a particular merchant. The system administrator could desire that if the daily collection of surveys for the merchant drops below 5 on a given day, an alert signal should be sent to the merchant or other party. This would allow the merchant to investigate the reason for the drop in response rate. The system 5

via event handler 32 or data monitoring system 30, for example, could

10 automatically send the merchant a signal such as an email to alert the merchant to the drop in rating. The merchant could then investigate the reason—perhaps competitors have dropped their prices, causing consumers to shop elsewhere, or perhaps the survey questionnaire is not being presented to purchasing consumers due to technical reasons. With respect to survey questionnaire

15 response rates, there could be correctable reasons for the alarm: a webmaster of a merchant website may have inadvertently deleted from a web page the prompt for a consumer survey questionnaire or placed it in a location on a web page that does not sufficiently attract consumer attention or interest. An alarm would give the system administrator an opportunity to investigate why the response rate has

20 decreased or increased allowing data flow to be maintained at desired or useful levels.

Alarm filters 28 may be associated with the processing modules 22 to give online merchants the ability to respond to changes in the marketplace at the earliest possible stage and to correct technical problems. In Fig. 3, the filters 28 are shown connected to one or more data channels 15 between server 14 and 5 database 16. An alarm filter 28 may be associated with any data channel or channels between two system components. The alarm filters 28 could also be associated with any other data channels 15 in system 5. The alarm filter sends a signal to an event handler 32 when a specified condition is met. The event handler 32 could be a computer system, a display device, telecommunications 10 device such as pager or phone, etc. The event handler 32 would inform a machine or person, such as a system administrator, end-user or other designated individual of the alarm.

Fig. 9 shows a flow chart of one possible alarm filter for monitoring survey questionnaire response rates. The alarm process 28.1 is started with a data run 15 in system 5. In decision step 28.2, the alarm system monitors a data channel for activity. If the system is not in an active data run, the alarm process ends in step 28.7. If the system is active, in step 28.3, synoptic response rates are calculated or input per a set interval of time. In step 28.4 topical response rates are calculated or input per a set interval of time. The calculated values are 20 compared with predetermined high thresholds in decision step 28.5. If a response is above a high threshold, a signal is sent to event handler 32. If the high threshold is not crossed, the values from step 28.4 are compared to a

predefined low threshold in step 28.6. If the low threshold is crossed, a signal is sent to event handler 28. If neither a high or low threshold is crossed, the system returns to the start step 28.1.

The software code for the foregoing may be implemented in any number  
5 of programming languages within the skill of persons in the art. As an example,  
MATLAB source code for one possible alarm filter process is as follows:

```
% 'rrmonitorHIGH.m' Response Rate Monitor computes response
% number thresholds for high (resps/day >= 30) arrival rates.
% Copyright Bizrate.com
10
clear all

% low arrival rate
M = 100; % length of data run (number of days)
15 lam = 500; % mean arrival rate (responses per day)
sqrtlam = sqrt(lam);
R = poissrnd(lam,1,M); % responses data array
Phigh = 0.9; Plow = 0.1; % high/low prob thresholds

20 nhigh0 = norminv(Phigh, lam, sqrtlam); % high number threshold
nlow0 = norminv(Plow, lam, sqrtlam); % low number threshold

N = 30; % length of data window for computing lam

25 figure(1)
clf
hold on
plot(R)
plot([0,M],[nhigh0,nhigh0],'r')
plot([0,M],[nlow0,nlow0],'r')
errorcount = 0;

30
for i = N:M-1 % cycle through the data run
    lastN = R(1,i-N+1:i);
    lam = sum(lastN)/N; % average of last N data points
    sqrtlam = sqrt(lam);
    nhigh = norminv(Phigh, lam, sqrtlam); % upper threshold
    nlow = norminv(Plow, lam, sqrtlam); % lower threshold
    plot(i+1,nhigh,'b.')
    plot(i+1,nlow,'b.')

40
    if R(i+1)>=nhigh % test for issuing 'high alert'
        if R(i+1)>= nhigh0
            plot(i+1,R(i+1),'go')
        else
            plot(i+1,R(i+1),'ro') % erroneous alert
            errorcount = errorcount + 1;
        end

45
    elseif R(i+1)<=nlow % test for issuing 'low alert'
        if R(i+1)<= nlow0
            plot(i+1,R(i+1),'go')
        else
```

```
plot(i+1,R(i+1),'ro')      % erroneous alert
errorcount = errorcount + 1;
end
end
5 end

grid
hold off

10 errorcount

% 'rrmonitorLOW.m' Response Rate Monitor computes response
% number thresholds for low (resps/day < 15) arrival rates.

15 clear all

% low arrival rate
M = 100;      % length of data run (number of days)
lam = 3;      % mean arrival rate (responses per day)
20 R = poissrnd(lam,1,M); % responses data array
Phigh = 0.9;  Plow = 0.1; % high/low prob thresholds

nhigh0 = poissinv(Phigh, lam)+1; % high number threshold
nlow0 = poissinv(Plow, lam);   % low number threshold
25 N = 15; % length of data window for computing lam

figure(1)
clf
30 hold on
plot(R)
plot([0,M],[nhigh0,nhigh0],'r')
plot([0,M],[nlow0,nlow0],'r')
errorcount = 0;

35 for i = N:M-1          % cycle through the data run
    lastN = R(1,i-N+1:i);
    lam = sum(lastN)/N;    % average of last N data points
    40 nhigh = poissinv(Phigh, lam)+1; % upper threshold
    nlow = poissinv(Plow, lam);        % lower threshold

    if R(i+1)>=nhigh % test for issuing 'high alert'
        if R(i+1)>= nhigh0
            plot(i+1,R(i+1),'go')
        45 else
            plot(i+1,R(i+1),'ro')      % erroneous alert
            errorcount = errorcount + 1;
        end

50    elseif R(i+1)<=nlow % test for issuing 'low alert'
        if R(i+1)<= nlow0
            plot(i+1,R(i+1),'go')
        else
            plot(i+1,R(i+1),'ro')      % erroneous alert
            errorcount = errorcount + 1;
        end
    end
55    end
end

60 grid
hold off

errorcount
```

65 Copyright in the foregoing code is claimed by BIZRATE.COM.

GANZLAW, PC  
PO Box 10105  
Portland, Oregon 97296  
Phone: (503) 224-2713  
Docket No.: BIZ/99-0008  
Express Mail No: EL326146496US Deposited March 29, 2000

### SATURATION LIMITED FORECAST MODEL

5

The present invention also provides a saturation limited forecast model (hereinafter "SLF Model") that forecasts the growth of a population from a set of early observations. This model may be integrated into system 5 as a processing module 22. The SLF Model may be used to predict various values of interest to 10 businesses. The forecasting methodology for e-commerce measures and consumer behaviors may be advantageously based on the saturation limited forecasting model of the present invention. For example, the model may be used to predict sales volumes for a category or categories of goods or services; number of females over the age of 18 participating in e-commerce; number of 15 merchants offering a certain category of goods or services. Values for e-commerce measures and consumer behaviors, such as the foregoing may be referred to herein as "e-commerce populations." An example of an SLF Model is more specifically shown in Fig. 10 as processing module 222.

More specifically, the SLF processing module uses available recent 20 historical data along with an estimated and/or available saturation "population" function as the basis for a differential equation that defines the growth of a "population" to a maximum attainable level. This differential equation actually embodies a family of realistic "penetration processes" that are found in nature and most areas of human activity. The results from this technical approach avoid

the errors often encountered in conventional "constant percent rate" predictor models.

The SLF Model embodies a non-linear mathematical programming solution to select the optimum set of process defining parameters from the 5 defined inputs. Once the nominal penetration function is obtained, the differential equation can be solved for any arbitrary time point in the future to provide the synoptic or nominal value forecast. Depending on the nature of the historical data set, the SLF Model also can extract the cyclic (i.e. seasonal or topical) variations for the predicted population. Additionally, by entering a range of 10 values for the inputs, the SLF Model will generate a "forecast funnel" that encompasses the confidence intervals of such input data.

In the SLF Model, the growth of a population may be forecasted from a set of early observations. To do this, we assume that there is a known or assumed, possibly time varying, population level  $P_0(t)$  or "pull function" which sets the 15 saturation limit to growth. A primary SLF Model operating principle is that growth is generated by an underlying process and that the rate of growth at any time is proportional to the unsaturated or remaining population. The proportionality is specified by a time varying penetration function  $r(t)$  which characterizes the total "level of effort" process. The population growth differential equation can then be 20 written as

$$\frac{dP}{dt} = r(t)[P_0(t) - P(t)]$$

A functional form, usually, but not limited to, a polynomial with coefficient vector  $\underline{c}$ , is then assumed for the penetration function giving  $r(\underline{c}, t)$ . The differential equation is then numerically integrated to generate the predicted population  $P(\underline{c}, t)$ . If we have a set of observations  $\{P_i | i = 1, \dots, N\}$ , then we can solve for the 5 value of  $\underline{c}^*$  that makes  $P(\underline{c}^*, t)$  a best fit to these points in, say, a least squares manner. In short, we numerically solve

$$\underline{c}^* = \arg \min_{\underline{c}} \left\{ \sum_{i=1}^N \left( P_i - \left[ \int_{t_0}^{t_F} r(\underline{c}, t) [P_0(t) - P(t)] dt \right]_{t_i} \right)^2 \right\}$$

10 where  $t_0$  may be specified (along with  $t_F$ ) or included in the optimized parameter set.

Turning now to the novel forecasting method in further detail, in [1] (bracketed numbers refer to the equations herein with the corresponding bracketed numbers) a population is penetrated at a rate proportional to a residual 15 population  $P_r$ . The proportionality is determined by a penetration function  $r(t)$  which captures the level of effort expended in order to penetrate  $P_r$ , and, possibly, maintain the achieved penetration  $P(t)$ . We further assume the influences of time-varying saturation level  $P_0(t)$  which yields the residual

$$P_r(t) = P_0(t) - P(t) \quad [1]$$

20 giving the penetration rate as

$$\begin{aligned}\frac{dP}{dt} &= r(t)P_r(t) \\ &= r(t)[P_0(t) - P(t)]\end{aligned}\quad [2]$$

For arbitrary  $r$  and  $P_0$  function, the solution to [2] is schematically shown in Fig. 11.

PROBLEM STATEMENT: Suppose we have a usable estimate of  $P_0$  and several values of  $P$ , say  $\{P_k; k=1, K\}$ , for arbitrary times  $t_k$  with  $K$  equal to the number of observed points. If we also assume a functional form for  $r(\cdot)$  that requires knowledge of a parameter vector  $\underline{c}^T = [c_1, c_2, \dots]$  of constants, then we would like to devise a method of using the  $P_k$  to extract  $\hat{r}(t) = r(\hat{\underline{c}}, t)$  such that [2] can be numerically solved to yield  $\hat{P}(t)$  and  $\sigma_p$ , which are graphically represented in Fig. 12. The  $\sigma_p$  bounds shown may be derived by stochastic sampling of the probability distributions characterizing the inputs or simply approximated from the solutions using the max and min ranges of the saturation function  $P_0$ .

Note the addition of  $t_0$  defined from  $P(t \leq t_0) = \emptyset$ . The inclusion of  $t_0$  as a hard constraint may only be required if we seek an analytical solution to [2] as discussed in [1]. In other words,  $t_0$  is implicit in the selection of expressing the  $t_k$ .

We need to extract the optimum  $\hat{\underline{c}}$  that yields  $P(\hat{\underline{c}}, t) = \hat{P}(t)$ . This optimum population growth function is derived from minimizing

$$J(c) = \sum_{k=1}^K [P(\underline{c}, t_k) - P_k]^2 \quad [3]$$

which then defines

$$\hat{\underline{c}} = \arg \min_{\underline{c}} J(\underline{c}) \quad [4]$$

5 The problem is that with any time-varying (i.e. realistic)  $P_0(t)$  it is not possible to solve  $\frac{dP}{dt} = r[P_0 - P]$  to yield an analytic expression as a function of some parameter vector such as  $\underline{c}$ . For a given  $\underline{c}$  we can obtain  $P(\underline{c}, t)$  only by numerical solution of the differential equation. This suggests the core computational scheme for  $\hat{P}(t)$  (SLF 222) shown in Fig. 10 in steps 222.2-222.16.

10 In SLF, the prime saturation or "pull" function drives  $\frac{dP_1}{dt} = r(t)[P_0 - P_1]$  with forecast horizon  $t_0, k_0$ . The prime or top level pull (saturation limit)  $P_0$  draws and meters the growth of  $P_1$ . These concepts are graphically shown in Fig. 13. Once  $P_1$  is generated, it can act as the pull function for the next contained/constrained population  $P_2$  and so on. Each saturation limited population  $P_l$  is generated from 15 an input set that contains its birth time  $t_{l,0}$  and a set of observed tuples  $\{P'_{l,k}, t_{l,k} \mid k = 1, K_l\}$ . (An example table of such tuples is presented in Fig. 14.)

Proper birth times are  $t_{l,0} \geq t_{l-1,0}$  for  $l=1, 2, \dots$ . Proper observations are such that

$$P'_{l,k} < P_{l-1}(t_{l,k}) \quad \forall l, k$$

This input horizon for all SLF forecasts is  $t_{0,K_0}$  since no successive pull function can be computed beyond that horizon. All birth times are such that  $t_{0,l} \geq t_{0,l-1} \forall l$ . Fig. 14 shows an example input array for SLF, including conjoined observations and predictions for all  $l$  and  $K_l$ .

5 The foregoing forecasting model may be implemented in standard programming languages for mathematical functions. Below is one example of the forecasting model coded in MATLAB programming language. The population being evaluated is a hypothetical Food & Wine commerce sector.

```
10 % Saturation Constrained Forecast Diff Eqn test solver
% Quadratic penetration, free birth time t0, weighted observations
clear all
15 global CONST OBS NOBS TINT W
% Prime Saturation ('Pull') Pop function constants
p = 16096; q = 199;
% Penetration Function constants r = a+bt+ct^2 and birth time.
20 a = .00001; b = 0; c = 0; t0 = 0;
CONST = [p q a b c t0];
% Master time interval
tmin = 0; tmax = 30;
TINT = [tmin tmax];
25 % observed pop levels
tobs = [9 10 11 12 13];
Pobs = [20 40 87 210 385];
30 W = [10 1 1 1 1]; % observation relative weights
%W = [1000 10 30 50 100]; % observation relative weights
W = W/sum(W); % normalized weights
35 % cyclic compensation inputs
nforecasts = 6; % nforecasts <= tmax - max(tobs)
Ncycle = 4; % length of cycle >= NOBS
fcomp = .8; % correction factor to compensations
40 NOBS = size(tobs,2);
OBS = [tobs; Pobs];
45 % compute Pen (J) Funct constants
X0 = [a b c t0]; % initial X value
ub = [.01 .01 .01 min(tobs)-1]; % upper bound on X
lb = [-.01 -.01 -.01 -10]; % lower bound on X
options = optimset('MaxFunEvals',2000);
[X,J] = fmincon('Jfun33',X0,[],...
[],[],[],lb,ub,[],options);
```

```
% compute the optimal P function from P(TINT(1))=0
[T,P] = ode45('dP33',[X(4) TINT(2)],0);
P0 = CONST(1) + CONST(2)*T;
5
% compute cyclically compensated forecasts

mindif = T - OBS(1,1);
maxdif = T - OBS(1,NOBS);
10 imin = find(abs(mindif) == min(abs(mindif)))-1;
imax = find(abs(maxdif) == min(abs(maxdif)))+1;
c = polyfit(T(imin:imax),P(imin:imax),3);
obsComps = OBS(2,:)./polyval(c,OBS(1,:))-1;

15 nstart = max(tobs)+1;
foreTimes = [nstart:nstart+nforecasts-1];
nforecasts = polyval(c,foreTimes);
compfore = zeros(1,nforecasts);
iOffset = mod(NOBS,Ncycle); % index offset in Obs array
20 for i = 1:nforecasts
    compfore(i) =
    nomforecasts(i)*(1+obsComps(iOffset+max(mod(i,Ncycle),(mod(i,Ncycle)==0)*Ncycle))*fcomp);
    %compfore(i) = nomforecasts(i);
end

25

% generate figure
figure(1)
clf
hold on
Forecasts = compfore; Measured = Pobs;
NomGrowth = P;
plot(tobs,Measured,'rx',foreTimes,Forecasts,'go')
plot(T,NomGrowth,'-',T,P0,'--')
35 axis([0 tmax 0 NomGrowth(foreTimes(size(Forecasts,2))))]
%axis([0 20 0 800])
set(gca, 'XTick', [0:1:tmax])
title('SLF Food&Wine')
xlabel('Qtrs - 95Q4=0')
40 ylabel('Online Sales $M')
legend('Measured','Forecasts','NomGrowth')
grid on
hold off

45
function J = Jfun33(X)
% criterion function for SALIMx, gjr 7apr99/27apr99

50 global CONST OBS NOBS TINT W

% solve DE and compute Saturation Level array
CONST(3) = X(1); CONST(4) = X(2);
CONST(5) = X(3); CONST(6) = X(4);

55 [T,P] = ode45('dP33',[X(4) 1.1*OBS(1,NOBS)],0);

% fit solution poly P(t) over observation interval

60 mindif = T - OBS(1,1);
maxdif = T - OBS(1,NOBS);
imin = find(abs(mindif) == min(abs(mindif)));
imax = find(abs(maxdif) == min(abs(maxdif)))+1;
c = polyfit(T(imin:imax),P(imin:imax),3);
Ppoly = polyval(c,OBS(1,:));
65
```

```
% compute criterion function
jcomps = Ppoly - OBS(2,:);
jcomps = (jcomps.*jcomps)*W';
J = sum(jcomps);

5

function dy = dP33(t,y)
10  global CONST OBS NOBS TINT W

P0 = CONST(1) + CONST(2)*t;

dy = (CONST(3)+CONST(4)*t+CONST(5)*t*t)*(P0 - y);
15
if dy<0
    dy = 0;
end

20 Copyright in the foregoing code is claimed by BIZRATE.COM.
```

25 DYNAMIC ACTIVITY ICON

The present invention also contemplates a processing module 22 that relates to the display of a dynamic icon that indicates to the user of a remote computer system some level of activity elsewhere in system 10. The dynamic icon could convey graphic or text-based information or both, as described in more detail below. In one possible embodiment, a dynamic icon or text is displayed through the web browser of a consumer's computer system or terminal that indicates to the consumer the level of activity by other visitors to a particular merchant's website. For example, the dynamic icon indicates the current/recent level of buying activity for the merchant. The activity could be indicated on a real-time basis or on defined time intervals. It could also relate to all transactions taking place at the merchant's site or on subcategories of transactions. In one embodiment, the dynamic icon is associated with ratings information about one

or more merchants whose ratings are made accessible to consumers through presentation server 26. The dynamic icon could also be associated directly with a merchant's own web server, via presentation server 26 of system 5, for viewing by consumers as they visit the merchant's website.

5        In the case of buying activity, the relevant data is received into system 5 as a survey questionnaire 100, for example. If survey questionnaires are served only to purchasing consumers, tracking the number of served questionnaires for a particular merchant, directly indicates the level of buying at a particular merchant. Alternatively, each survey questionnaire returned by a consumer  
10      could be correlated to buying activity (A correlation is necessary because not all purchasing consumers will complete and return a survey questionnaire.) The survey questionnaire contents may also be used to determine the activity level of other matters contained or associated with a survey. For example, the survey questionnaire may ask the purchaser to describe items purchase, quantities,  
15      prices, etc.

          Alternatively, electronic cookies may be used to follow consumers to certain categories of websites or to specific websites. The cookies could be received by data capture server 14. The presentation server could let consumers and/or merchants know which websites are receiving the most traffic. The traffic  
20      could be reported, broken down by category of trade. For example, the online music store website with the most traffic. This should help consumers identify where there are special promotions or available products, etc. This should help

merchants understand what their competitors are doing for market research purposes or for competitive response.

The methodology for a dynamic icon that communicates the level of activity of some predefined subject matter is well within the skill of persons in the

5 art. For example, Fig. 15 shows a flow chart for an activity module process 23 that receives data input from a data source, such as database 20. The example is in terms of the level of sale transactions on a point-of-sale website. In step 23.1, the point of sale ("POS") data for a merchant is read from database 20. In step 23.2, the number of POS transactions for the merchant is calculated. In 10 step 23.3, the calculated value is returned to display database 24. The calculated value is then accessible to presentation server 26 in step 23.4.

Presentation server 26 presents the calculated value as a dynamic icon accessible to remote computer systems through their web browser, for example.

In another embodiment, an activity-level applet is received and stored by a 15 consumer's computer. This is reflected in Fig. 15 at step 23.5. The applet generates a dynamic icon in the form of a flashing dot, for example. The icon flashes in proportion to the level of activity on a particular merchant site. The consumer's computer is informed of the activity level by accessing a data source that communicates data representative of the activity level. This is reflected in 20 step 23.6 of Fig. 15. The data is input into the activity-level applet that creates an output in the dot flashing at a rate proportionate to the level of buying at the merchant's site.

The present invention contemplates that a plurality of dynamic icons may be associated or arrayed with each other. By simultaneously displaying more than one dynamic icon, one or more icons may serve as reference points of activity relative to one or more other icons. In one embodiment, at least two icons

5 are simultaneously displayed, each icon representing activity levels for a different merchant. In a preferred embodiment, the merchants provide competitive goods or services. In another possible embodiment, the dynamic icons are displayed on a consumer's computer and are associated with hyperlinks to the merchants' websites.

10 In a further embodiment, there are at least two icons simultaneously displayed, one icon representing activity at a merchant website, and the other icon being a reference icon that represents a predefined level of activity against which other displayed dynamic icons may gauged. In a further embodiment, there are at least three icons simultaneously displayed, at least two of which

15 represent activity at competitive merchants, and at least one of which represents a reference icon. In addition, the present invention also contemplates simultaneous display of multiple groups of competitive merchants. It also contemplates that the activity level may represent a variety of things, and is not limited to level of point-of-sale activity. For example, it could indicate level of

20 traffic at the site; types of products or services being sold or offered; numbers of a product or service type or category being sold; inventory levels; special promotions or discounts; buyer demographical attributes, etc.

Consumers could investigate websites based on their interest in a level of indicated by a dynamic icon. For example, a relatively high rate of activity could mean exceptional prices, special offerings, good reputation, etc. This gives consumers (or merchants) a basis for comparing the level of activity between 5 merchants. It also helps promote consumer confidence in e-commerce.

With respect to the foregoing embodiments, a Poisson random number generator may be used to generate a sequence of numbers that represents arrival rates. The number of survey questionnaires served or the number of completed surveys received, for a predetermined interval, could be the basis for 10 the input of the mean rate into the Poisson random number generator. The mean arrival rate would then be sent from a presentation server or communicated to a consumer's computer, for example, as input into the activity-level applet implementing the Poisson random number generator on a consumer's computer.

15 The dynamic icon is not limited to a flashing dot. It could be any number of things, including a numerical value; textual description of activity; a graphic image that connotes a level of activity, such as a flashing colored, say, red dot; an audio-based icon that verbally or connotatively describes some level of activity, etc.

20

### INFOMEDIARY SERVICE

The foregoing embodiments, including DS, SLF, response rate alarm filters, and dynamic icon may be incorporated into a system that assists consumers in making e-commerce decisions, particularly which merchants to purchase from. (The system would also be applicable and readily adaptable to 5 business-to-business e-commerce).

In one possible embodiment, the presentation server 26 is a web server that hosts an "infomediary website" (as described by Hagel, John III and Singer, Marc in *Net Worth*; Boston, Mass; Harvard Business School Press 1999) that consumers may access through computer systems via the Internet or other 10 network. The website includes web pages having a list or table of merchant websites. The list or table could categorize merchants in any way, including by sector, by size, by how long they have been in engaged in e-commerce, by geography, by customer satisfaction ratings, by specials offered, etc. A displayed list of one or more merchants may be displayed to a consumer by the 15 consumer selecting desired categories from a menu on a web page.

Alternatively, the displayed list may be generated based on search criteria entered by a consumer and presentation of results by a search engine. Each merchant displayed could be associated with a dynamic icon showing selling activity, for example.

20 The consumer may access a displayed merchant's website through hyperlinks in the list or table. Preferably, the list or table includes ratings for each displayed website, to help consumers evaluate the website for relevant attributes,

such as those of survey questionnaires 100 and 200. Other information about the merchant could also be displayed or accessed through associated hyperlinks.

Ratings and other information could be based on data processed through the DS, SLF, or any other module of the present invention. The list or table also may

- 5 include icons or information denoting merchant websites that participate in consumer rebate programs. The infomediary website may also be linked a merchant's website for access to data about availability of products or services, pricing, etc. The foregoing has been in terms of a web page showing a listing or table of merchants. The foregoing discussion would also apply to a web page
- 10 with a listing of goods or services by category. Listed goods or services products could be associated with merchants selling the desired goods or services.

Still further, proximity searching means may be presented to the consumer for finding nearest neighbors to a displayed item on the infomediary website relative to specified attributes of an item. In this regard, US Patent No.

- 15 5,983,220, entitled "Supporting Intuitive Decision In Complex Multi-Attribute Domains Using Fuzzy, Hierarchical Expert Models", commonly owned by the assignee of the present invention, is hereby expressly incorporated by reference for all that is taught therein.

- 20 A membership program may be used on the infomediary website, allowing registered members preferred access to information, automated notifications about subject matter within their scope of interest, or other preferential treatment.

A registered member would be someone providing predetermined categories of

information called for by the infomediary website, such as personal data, shopping preferences, participation fees, etc. The infomediary website could notify a registered member of discounted prices for goods or service in which they have expressed an interest. The infomediary website could also arrange for

5 incentive programs for its members such as special discounts, rebates, frequent buyer programs, etc. could be offered by or through the infomediary. For example, registered members purchasing from a participating merchant could receive a rebate according to predefined terms of a rebate program.

Figs.17-23 show web pages illustrating features in accordance with the

10 foregoing discussion. The web pages could be served by presentation server 26 in earlier Figures. Fig. 16 is an example of a home page for an infomediary web site. The page describes a number of products and services that may be browsed or searched by a consumer visiting the web site. The page also announces incentive programs and special offer programs. The names, 15 descriptions, and symbols on the web pages of Figs. 17 to 23 may be presented in the form of links to other pages inside or outside the infomediary website. The web page also has means to allow a consumer to search for products or services via categories listed on a pull down menu. A search prompt is presented by selecting a category from the pull down menu and clicking on the go button. The 20 web page of Fig. 16 also presents certain popular products in the left margin of the page.

Fig. 17 is a web page from the infomediator web site of Fig. 16. This page represents the "Computer Hardware & Software" category from the page of Fig. 16, and may be reached by clicking on that category name in the web page of Fig. 16. The page lists subcategories of products. The page also includes a 5 search prompt for searching by product identifiers, such as SKU numbers. There are additional pull-down menus to produce search prompts to search or find other categories or merchants. Incentive programs and special offers may also be presented or accessed through the web page.

Fig. 18 is a web page from the infomediator web site of Fig. 16. This page 10 represents the "PDAs" category shown on the Computer Hardware & Software category web page of Fig. 17. The "PDAs" web page of Fig. 18 may be reached by clicking on that category name on web page in Fig. 17. The Web page of Fig. 18 lists merchants carrying PDA products. The merchant names are associated with overall ratings in the form of 1-5 stars. The ratings may also be 15 distinguished as to merchants that participate in a customer ratings program. A consumer visiting the infomediator site may learn details of such a program by a linked web page describing the program, e.g. the web page of Figs. 19a-b. Merchants that allow post-purchase surveys of consumers are indicated with gold rating stars and non-participating merchants with silver stars, as described 20 in Figs. 19a-b.

A consumer can learn more about a merchant listed on the web page of Fig. 18 by clicking on the merchants' name. The merchant name may, for

example, link to further ratings information or provide comments or reviews by consumers or others. The web pages with such information may be on or off the website of the infomediary. Special offers such as rebates are also associated with a listed merchant, as seen in Fig. 18. An "on time" rating attribute is also

5 associated with a listed merchant. The ratings and other information presented on the web page may be derived from one or more of the processing modules of discussed above, including the DS and SLF processing modules. Other attributes, such as described for surveys 100 and 200 of the earlier Figures may be associated with a listed merchant. As indicated in Fig. 18, searching and

10 sorting features may also be included on the web page to search or sort for desired rating attributes, product pricing, product characteristics, product availability, etc.

A dynamic activity-level icon, as discussed above may also be associated with merchants listed on the web page of Fig. 18. In this case it is a dot of

15 varying size, the larger the dot the higher the activity level. Consumers may select to browse or shop at a merchant's website by clicking on the "GO SHOP" icon seen of the web page. The icon provides a direct link to the merchant's website. Preferably the link is to a web page on the merchant's website that includes the product of interest. Fig. 20 represents the web page of the

20 merchant named "eCost.com". The web page shows price and product information for a PDA.

The web page of Fig. 18 also lists in the right hand margin a number of PDA products considered to be top picks according to selected criteria believed to be relevant to consumers. Clicking on a product may take the consumer to a list of one or more merchants offering the product.

5 The web page of Fig. 18 page also includes search and sort prompts for searching by merchant ratings, product identifiers or characteristics, such as manufacturer or keywords. There are additional pull down menus to produce search prompts to search or find other categories or merchants. Incentive programs and special offers may also may be presented or accessed through the

10 web page.

An example referral incentive program offered on the infomediary web site is described on the web page of Fig. 21.

An example of special offers presented on the infomediary web site is described on the web page of Fig. 22.

15 An example form for becoming a registered member of the infomediary is presented on the web page of Fig. 23.

The foregoing embodiments are for illustrative purposes and are not intended to be limiting, persons skilled in the art capable of appreciating other embodiments from the scope and spirit of the foregoing teachings.